

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

EMMA C., et al.,
Plaintiffs,
v.
TONY THURMOND, et al.,
Defendants.

Case No. [96-cv-04179-VC](#)

**ORDER RE STATE'S COMPLIANCE
AT PHASE 2**

After inheriting this consent decree from the previously assigned judge, the Court set up a process for examining whether the State of California does a legally adequate job of monitoring school districts for compliance with the Individuals with Disabilities Education Act (IDEA). The process has four phases, the first of which is complete. That phase addressed the state's data collection activities, and the Court held that the state will – after correcting a couple isolated but significant defects – be deemed in compliance with its federal statutory obligations in this area.

Thus began the second phase, which examines how the state analyzes the data it collects to determine which districts are in need of intervention, and to determine what type of intervention is called for. The matters addressed in this phase are more central to the state's responsibility to monitor school districts to ensure that disabled children are receiving an appropriate education as required by the IDEA.

Unfortunately, California's system for flagging school districts for intervention is riddled with serious defects. To give just three examples:

- The state attaches the classification of “meets requirements” to hundreds of small school districts despite conducting virtually no analysis of those districts.

- The state takes minimal steps to flag problems with the delivery of services to preschool-aged children with disabilities, even while acknowledging the importance of early intervention to IDEA compliance.
- The state’s system for identifying the most troubled districts appears irrational, resulting in the selection of districts for intensive intervention that are less in need of it than the ones passed over.

It appears, at least so far, that current leaders in the Special Education Division of the California Department of Education don’t deserve much of the blame for this. They inherited many of the problems, candidly acknowledge most of them, and are committed to improving the state’s monitoring activities. They have been forthcoming and cooperative, even while their lawyers from the Attorney General’s Office have sometimes been nonresponsive or obstructionist. Accordingly, for now, there is no reason to seriously consider holding the state in contempt for failure to comply with the consent decree. Nonetheless, the defects in the current system are so serious, and so numerous, that they significantly interfere with California’s ability to monitor how school districts are serving disabled children. The state cannot move on to Phase 3 until it addresses – or shows that it’s well on its way to addressing – these problems.

I. BACKGROUND

A.

The IDEA requires states that accept federal assistance to provide a free and appropriate education to all children with disabilities. 20 U.S.C. § 1412(a)(1)(A). As a practical matter, this responsibility falls largely on individual school districts. In recognition of this, the statute requires states, in turn, to conduct effective oversight of the school districts. *Id.* § 1416(a)(1)(C). The consent decree in this case requires the state to demonstrate that it has an adequate oversight system in place.¹

¹ As discussed in the May 18, 2018 order, the precise question is whether the state has a monitoring and enforcement system that is adequate with respect to Ravenswood City School District. The state has explained, however, that it intends to establish compliance with the consent decree with reference to its normal statewide system for monitoring school districts,

This Court's May 18 order established a four-phase monitoring process for assessing the legal adequacy of the state's oversight system. Dkt. No. 2387 at 1–2. Phase 1 involved scrutiny of the data the state collects from school districts. Specifically, the inquiry was whether the state collects enough data, and the right types of data, to enable it to effectively monitor districts. The Court concluded that although isolated legal deficiencies in the state's data collection system must be addressed, the state was largely compliant in this area. The Court thus concluded that the state could move on to Phase 2 and could cure the isolated data collection deficiencies in the subsequent phases.

Phase 2 – the current phase – examines how the state analyzes the data it collects. To satisfy its monitoring obligations under the IDEA, the state must do an adequate job of flagging the districts that are failing, or struggling, to provide an appropriate education to students with disabilities. And it must do an adequate job of deciding what type of intervention is necessary in a given district. Phase 3 will examine how the state, after it has decided which districts to select for further monitoring, actually executes that monitoring. Phase 4 will examine the state's written policies and directions governing school district compliance with the IDEA.

Phase 2 essentially poses two questions. The first is whether the state translates the data it collects into metrics capable of identifying school districts that may be falling short on their obligation to provide an appropriate education to students with disabilities. The second question is whether the state's methods for sorting districts for inclusion in (and exclusion from) its various monitoring activities is adequate.

In addition, the IDEA's implementing regulations require states to issue annual compliance determinations for school districts. 34 C.F.R. § 300.600(a)(2). The state's process for selecting districts for monitoring also determines whether the districts are classified by the state as complying with federal law. School districts that are selected for monitoring are classified as

rather than developing a narrower remedy that would apply to Ravenswood only. *See* Dkt. No. 2387 at 1. Because Ravenswood primarily serves preschool students and students in grades kindergarten through eighth grade, this order does not consider elements of the state's monitoring program that would apply, for example, exclusively to high school students.

“needs assistance,” “needs intervention,” or “needs substantial intervention” if problems persist. States must prohibit these school districts from reducing their “maintenance of effort,” meaning their allocation of non-federal funds for special education services. 34 C.F.R. § 300.608(a). School districts that the state decides are not in need of any monitoring are typically classified as “meets requirements.” Phase 2 therefore inherently includes an inquiry into whether the state’s sorting methods adequately ensure that districts labeled as “meets requirements” do not suffer serious deficiencies in serving disabled children.

Phase 2 took roughly the same format as Phase 1. The Court received written submissions from the parties, a report from the court monitor outlining his conclusions, and an amicus brief from the Morgan Hill Concerned Parents Association. The written submissions were followed by two days of evidentiary hearings. Kristen Wright, the Director of the Special Education Division for the California Department of Education, Shiyloh Duncan-Bercerril, the Division’s Education Administrator, and Alison Greenwood, the Division’s Quality Assurance Administrator, testified during Phase 1 and returned to testify for Phase 2. After the hearings, the Court ordered the monitor to conduct supplemental data analyses and ordered the state to provide all data necessary for those analyses. The monitor and his data consultant, Dr. Susan Wagner, presented their conclusions at a third evidentiary hearing, and state policymakers offered further testimony in response.² The monitor then conducted a final set of data analyses.

B.

As explained in previous orders, the purpose of this oversight is to ensure that the state complies with federal law. The IDEA and its implementing regulations offer only general guidance for what states must do to satisfy their monitoring and enforcement obligations. The state’s monitoring activities must focus on “improving educational results and functional outcomes for all children with disabilities,” and ensuring that states meet the IDEA’s requirements, with a special focus on “priority areas” enumerated in the statute. 20 U.S.C.

² Alison Greenwood did not testify at the third evidentiary hearing. Stacy Wedin, a policy consultant in the state’s Special Education Division, testified in her place.

§ 1416(a)(2), (a)(3). The state must use “quantifiable indicators and such qualitative indicators as are needed to adequately measure performance in the priority areas.” 34 C.F.R. § 300.600(c). When a district appears from the data to fall short on its obligations, the state must respond with an appropriate enforcement action to correct the noncompliance “as soon as possible.” 34 C.F.R. § 300.600(e). The IDEA, however, does not require states to adopt any particular approach for monitoring. By specifying the ends, but leaving the means to the states, the IDEA strikes a balance between federal authority and the states’ historic discretion in the design and control of their education systems.

For purposes of this consent decree, however, the standard for compliance is relatively straightforward. The state will not be found to be out of compliance simply because the plaintiffs or the court monitor have identified isolated deficiencies or ways in which the monitoring system could become more effective. But if the state’s chosen procedures are so deficient that they significantly hinder its ability to monitor school districts, the state will not be found compliant merely because the statute does not expressly forbid those choices.

This standard has important implications for this phase. Any evaluation of the state’s data analysis activities must pay close attention to how that data is actually used. Because certain problems may occur in tandem – for example, districts that frequently suspend students with disabilities may also have poor performance on statewide assessments – different metrics may yield similar information. Therefore, an imprecise metric may not compromise the overall monitoring system. Further, the importance of precision may turn on the importance of the metric. The state need not use the most granular data possible if the data is less important in the context of the overall system, or if the state has set targets for that data that sufficiently ensure that any poor performance will lead to monitoring.

These realities underscore the importance of reviewing the state’s data analysis activities as a whole. No monitoring system is perfect. Identifying a theoretical concern with an individual metric is not sufficient to find the state out of compliance with federal law. It must be shown that the concern, either alone or in combination with other issues, significantly interferes with the

state's ability to evaluate school districts and to identify the ones in need of intervention.

C.

As discussed in detail in the Phase 1 order, during the school year, school districts submit large swaths of data about all students, with and without disabilities, to databases maintained by the state. *See* Dkt. No. 2428 at 3. This process constitutes the “first tier” of the state’s monitoring system.

The state then translates those data into metrics that align with key directives of the IDEA. For example, one directive is that schools educate students with disabilities in the “least restrictive environment.” 20 U.S.C. § 1412(a)(5). In general, this means that students with disabilities must be taught in general education classrooms, alongside their nondisabled peers, as often as reasonably possible. Schools may only remove students with disabilities from general education classrooms when the nature or severity of the child’s disability requires it. To evaluate district performance in this area, the state uses three different metrics. For school-age children, the state calculates the percentages of disabled students who are taught (i) in general education classrooms for greater than 80% of the day; (ii) in general education classrooms for less than 40% of the day; and (iii) in separate placements. For preschool-aged children, the state looks to the percentages of disabled students in (i) general early-childhood programs; and (ii) separate placements. Dkt. No. 2455-1 at 11–12, 14. The state compares school districts’ performance on each metric to a target that policymakers set in consultation with various groups, including local administrators, parent groups, and advocacy groups. Dkt. No. 2455-1 at 4.

In other areas, the state uses a system called the “Dashboard.” The Dashboard is a visual depiction of district performance on a grid. The vertical axis of the grid represents the school district’s “status,” or the district’s current performance. The horizontal axis on the grid represents the degree to which the school district improved or regressed from the previous year. Combining information from both axes results in a color: red, orange, yellow, green, or blue. An example from the state’s written submissions is included here:

Performance Level	Declined Significantly from Prior Year (by more than 15 points)	Declined from Prior Year (by 3 to 15 points)	Maintained from Prior Year (declined by less than 3 points or increased by less than 3 points)	Increased from Prior Year (by 3 to less than 15 points)	Increased Significantly from Prior Year (by 15 points or more)
Very High +45 points or higher in Current Year	Green	Green	Blue	Blue	Blue
High +10 to +44.9 points in Current Year	Green	Green	Green	Green	Blue
Medium -5 points to +9.9 points in Current Year	Yellow	Yellow	Yellow	Green	Green
Low -5.1 to -70 points in Current Year	Orange	Orange	Orange	Yellow	Yellow
Very Low -70.1 points or lower in Current Year	Red	Red	Red	Orange	Orange

Dkt. No. 2455-1 at 24.

This is the Dashboard that the state uses to evaluate disabled students’ performance on statewide assessments. The state uses a metric called “distance from standard.” *See* Dkt. No. 2455-1 at 22–23. For each student, the state calculates the distance between the student’s score and the score needed to establish that the student met the academic standards relevant to that assessment. The state then calculates the average “distance from standard” for each school district and reduces the score if the district failed to ensure that enough disabled students actually took the assessment. The circled squares show that districts with a distance from standard of 10–44.5 points that improved their performance between 3–14 points from the prior year will receive a “green.” Districts that receive a “red” or “orange” are treated as missing the target for this metric. In addition to performance on statewide assessments, the state currently uses the Dashboard to evaluate districts’ suspension practices.

D.

After the data are translated into a set of metrics and compared to targets, the next step is to determine whether a district should be chosen for further monitoring, and if so, what type of monitoring activity to conduct. The degree to which the state effectively executes its monitoring

activities will be explored in depth during Phase 3. But evaluating how the state decides which monitoring activities to conduct, and for which districts, requires at least some understanding of the substance of those monitoring activities and how they differ from one another.

In general, a district that performs poorly on any individual metric will be selected for a type of targeted monitoring. A district that performs poorly across many different metrics, or in certain instances, the same metrics for several consecutive years may be selected for more intensive monitoring. In all, the state currently performs five types of monitoring relevant to this case – three targeted monitoring activities that relate to a specific area of poor performance by a district, and two intensive monitoring activities that involve greater intervention.³

The state has dubbed the three targeted monitoring activities “performance indicator review,” “data identified noncompliance review,” and “disproportionality review.” A district is selected for performance indicator review if it fails to meet a target for a particular metric that the state believes is closely tied to student outcomes. Dkt. No. 2506 at 6. For example, a district that receives a red or orange on the Dashboard for poor performance on statewide assessments by disabled students, or for excessive suspension of disabled students, will be selected for performance indicator review in the pertinent area. A district selected for review must perform a “root cause” analysis into the reasons for its inadequate performance in that area and must submit an improvement plan to the state. Dkt. No. 2501 at 7–9. The state does not formally

³ The monitoring activities discussed in this order exist alongside other accountability measures developed by the state. For example, state officials testified during the hearings that the California Department of Education may separately offer school districts “differentiated assistance” based on some of the same performance criteria that it uses in determining whether that district is providing an adequate education to disabled students. Dkt. No. 2506 at 120–21. In addition, a significant portion of school districts receiving differentiated assistance do so specifically because of their outcomes for students with disabilities. Dkt. 2478-1 at 4. Because differentiated assistance is not administered by the Special Education Division of the Department of Education, which is the division responsible for IDEA compliance and the division whose activities are being scrutinized here, the state provided little information about it in its written submissions. Further, the information presented at the evidentiary hearings does not explain in detail the forms of assistance that these districts may receive, and leaves it unclear whether those forms of assistance may assist the state and school districts in complying with the IDEA. If the state believes that a fuller picture of differentiated assistance would assist the Court in evaluating the state’s monitoring systems, then it should provide more information in future submissions for Phase 2 and/or Phase 3, when the inquiry turns to the content of the state’s monitoring activities.

supervise the district's implementation of the plan but looks to see whether the targets are met for the following year. If the district remains in performance indicator review for multiple years, it must perform a "record review," meaning it must select ten individual student records and review them for any patterns related to the noncompliance. Dkt. No. 2506 at 161–63.

The second targeted monitoring activity, which has the unfortunate name of data identified noncompliance review, addresses poor performance on metrics that are related to timeliness. For example, before a child may receive special education services, the district must conduct an initial evaluation to determine whether the child has a disability. The IDEA requires districts to conduct these evaluations within 60 days after receiving parental consent. 20 U.S.C. § 1414(a)(1)(C)(i)(I). The state flags districts for monitoring whenever they fail to meet this and any other statutory deadline for any student. The district must correct all identified noncompliances (for example, by holding any past-due initial evaluations). The district must also submit a root cause analysis that explains why the district has missed the deadline (or deadlines) and must submit data to the state demonstrating that it has corrected the problem. If the district continues to miss deadlines after that, the state may order it to take further corrective actions, and may even withhold funding to the district. Dkt. No. 2506 at 24.

The third form of targeted monitoring, disproportionality review, relates to the IDEA's requirement that states prevent discrimination against children on the basis of race or ethnicity. 20 U.S.C. § 1418(d). To use one concrete example, school districts must take care to avoid suspending Hispanic disabled children at a disproportionately higher rate than all disabled children who are not Hispanic. *See* 34 C.F.R. § 300.646(a)(3). The state is given the discretion to determine how marked the discrepancy must be before districts become "significantly disproportionate." 34 C.F.R. § 300.647(a)(7). Once selected, the state evaluates the districts' policies and procedures and a sample of student records for compliance with federal law (for example, to ensure that all IEP meetings were conducted with a general education teacher present). Dkt. No. 2455-1 at 54; Dkt. No. 2506 at 25–26. The district must correct all noncompliance identified by the state.

The state's two intensive monitoring activities are comprehensive review and significant disproportionality review. Comprehensive review is reserved for districts experiencing the most serious performance issues. The decision to select districts for comprehensive review is based on a total score derived from performance on many different metrics. Poor performance on any individual metric will, in addition to potentially leading to selection for targeted review, cause the district to lose points from its total score. Dkt. No. 2469 at 57. After all districts are scored, the state sets a cut score and selects all districts falling below that score for comprehensive review. In contrast to performance indicator review and data identified noncompliance review – which largely involve self-analysis by the district – the state puts its own boots on the ground during comprehensive review. The state, rather than the district, develops the monitoring plan after analyzing the district's compliance history. It will interview parents and administrators and conduct staff trainings as needed. A district selected for comprehensive review may remain in monitoring for several years. Dkt. No. 2506 at 30–34.

The state's other intensive monitoring activity, significant disproportionality review, applies to districts that California has determined to be “significantly disproportionate” in any area for three consecutive years. During this monitoring activity, the state takes an even harder look at the district's policies and practices. In addition to the activities performed during targeted disproportionality review, districts in intensive monitoring must develop an improvement plan that the state dubs an “early intervening services plan,” which must be approved by the California Department of Education. The state also makes various forms of technical assistance available to districts, both through the state directly and through contractors, to assist the district in developing their plan and getting back into compliance. Districts must set aside 15% of their IDEA-based funds to finance the implementation of the plan and must report on their progress to the state. Dkt. No. 2506 at 28–30.

II. OBVIOUS DEFECTS IN THE STATE'S DATA ANALYSIS SYSTEM

The state's process of determining which school districts need further monitoring, as well as its process for deciding how intensive that monitoring should be, is riddled with defects. Some

of these defects are small and easily fixable, but several are so fundamental, or so obviously contrary to the IDEA (or both), that the state cannot get out from under this consent decree without fixing them. Furthermore, in contrast to Phase 1— where the state largely established compliance, such that it made sense to move on to Phase 2 of the monitoring process concurrent with the state’s efforts to fix a couple of outstanding legal defects in its data collection program — the flaws in the state’s system for identifying districts in need of intervention are so severe that they preclude a transition to the next phase. Before moving on to Phase 3, the state will be required to demonstrate that it has meaningfully addressed these problems (or at least that it’s well into the process of addressing them).

A.

As a preliminary matter, there are several areas where the state essentially conducts no assessment of school district performance at all, even though the IDEA requires it. The policymakers testified at the hearings that they are currently developing data analysis procedures for almost all of these areas. But there is nothing in place now.

Small school districts. The state requires school districts to have a minimum number of disabled students (or “n-size”) before assessing those districts with the Dashboard and with most other metrics. *See* Dkt. No. 2478-3 at 2. Almost all of the state’s metrics (including all metrics for academic performance, least restrictive environment, discipline, and disproportionality) have an associated minimum n-size. The only real exception is for the metrics relating to timeliness. Thus, the state is conducting no actual substantive evaluation of small districts. For example, the state did not evaluate performance on English and Language Arts and Mathematics assessments for school districts serving approximately 60,000 disabled students. It did not examine suspension practices for districts serving approximately 40,000 disabled students. Dkt. No. 2494, Exhibit 1 at 6.

Because these small districts are not being assessed on many metrics, they are functionally exempt from many targeted and intensive monitoring activities. During the hearings, state officials testified that over 1,000 school districts — comprising roughly 600 charter schools

(which the state treats as separate school districts for its data analysis activities) and roughly 400 small school districts – had fewer than thirty disabled students. Dkt. No. 2507 at 36; Dkt. No. 2494, Exhibit 1 at 6. Consequently, these districts could not be selected for targeted review for academic assessments or for their suspension rates and could not be evaluated for comprehensive review based on their performance on those metrics. *See* Dkt. No. 2478-3 at 2.

Moreover, it bears recalling that the state classifies any district not selected for monitoring activities as “meets requirements” for the purposes of the state’s federal reporting obligations. Therefore, these small districts are particularly likely to be classified by the state as compliant with the IDEA irrespective of their actual performance. Of the 499 school districts not selected for any monitoring process and determined by the state to meet requirements, only 132 had a census count of greater than 30 students with disabilities. Dkt. No. 2516 at 8. In other words, the state is telling the federal government that hundreds of school districts “meet requirements” under the IDEA even though the state has effectively exempted them from monitoring.

To be clear, the state has identified sound reasons for excluding small districts from the data analyses it ordinarily performs. A small number of disabled students in a given district may cause that district’s metrics to vacillate wildly based on the performance of just a small handful of students (or even just one student). Dkt. No. 2506 at 202. But this cannot account for the state’s failure to develop any alternative protocol for assessing these districts, given that the IDEA requires states to monitor all districts. The policymakers acknowledged this at the hearings, and they testified that they have already begun researching methods of closing this gap, including by aggregating small districts – perhaps with similar features or within the same county – into larger units for data analysis purposes. There may be several different ways to adequately include these districts, and how best to do so is a question for the policymakers. For now, it’s enough to say that the state’s failure to include them is a failure to comply with its monitoring obligations under the IDEA.

Preschool review. There are roughly 90,000 preschool-aged children with disabilities in

California. Dkt. No. 2478-1 at 23. A district that does poorly at placing preschool-aged disabled children in regular early-childhood programs may be selected for targeted performance indicator review. If that district does not put individualized education programs (“IEPs”) in place in a timely manner for very young children transitioning into early-childhood programs from early intervention programs (the “Part C to Part B” transition), or fails to conduct initial evaluations on time, that district will be selected for data identified noncompliance review. Finally, a district that performs very poorly with respect to preschoolers and school-aged children may be selected for comprehensive review, although, as discussed in the next section, the state’s selection formula makes this somewhat unlikely. This represents the full extent of the state’s monitoring for preschool-aged children.

State policymakers candidly testified that a separate preschool-related monitoring activity is necessary to close this gap, because preschools face unique issues that cannot be adequately addressed through other monitoring activities. Dkt. No. 2506 at 93–94. The state is currently in the process of designing such a system, on similar lines as comprehensive review. In addition to looking at districts’ preschool placements and timeliness in carrying out the Part C to Part B transition, the state would consider districts’ suspension practices, participation on preschool assessments, and performance on those assessments in selecting districts for review.

In its written submissions, the state described a selection methodology for identifying preschools for intensive monitoring, and described a set of monitoring activities, which at times made it seem as if this system were already in place. *See, e.g.*, Dkt. No. 2455-1 at 10 (“[T]he State addresses [least restrictive environment] in a number of monitoring activities, including Performance Indicator Review, Preschool Monitoring ... and Comprehensive Review.”), 52 (“The Preschool Review examines issues related to placement, suspension, child find, and the provision of FAPE for students 3 through 5 years of age.”); Dkt. No. 2478-1 at 8 n.16 (“[The state] identified . . . 39 [school districts] for Preschool Review.”). It did not become apparent until the hearings that this monitoring activity does not yet exist, in any form. As with the small school districts, the state’s failure to adequately monitor preschool children puts it out of

compliance with federal law.⁴

Mediation. The IDEA requires school districts to establish mediation procedures. 20 U.S.C. § 1415(e). In mediation, an impartial party helps to resolve disputes between parents and the school district regarding the district's obligation to provide special education services. The IDEA requires states to pay for the costs of mediation and to prioritize mediation in their monitoring activities. 34 C.F.R. §§ 300.506(b)(4), 300.600(d)(2). But the state does not assess districts' mediation practices. This is true even though the state collects data regarding districts' mediation practices and reports the percentage of mediations resulting in successful agreements to the federal government. Dkt. No. 2390-1 at 40–41.

The state defends this by noting that mediation is a voluntary process that requires the consent of both the parents and the school district, which means that the state cannot force recalcitrant districts to mediate. Even assuming that's true on a case-by-case basis, it's far from obvious that the state couldn't take action against districts that refuse to use mediation as a blanket matter (or, for that matter, seriously neglect this tool). Nor does it follow from the voluntary nature of mediation that the state need not analyze whether districts are doing a good enough job of using it. As the drafters of the IDEA have signaled, mediation is a useful alternative for parents who lack the resources to hire a lawyer to assist with a due process complaint, or who prefer more informal and nonadversarial methods for resolving disputes. Analyzing mediation data could therefore reveal trends not separately captured by an analysis of formal due process complaints or complaints made to the state. The state could also combine mediation data with other data – for example, districts' suspension rates or assessment outcomes – to form a more complete picture of district performance.

To be fair, the state's failure to assess districts for their use of mediation is less significant than some of the other problems discussed in this section. There is even an argument

⁴ It would be premature to make any formal conclusion regarding the state's anticipated selection methodology before any monitoring system is actually in place, but it's worth noting that the state's proposed formula closely mirrors the formula used for comprehensive review, which, as discussed in subsection II.B, contains serious flaws.

that the failure to analyze mediation data does not, on its own, significantly interfere with the state's ability to fulfill its monitoring obligations under the IDEA. But in contrast to some of the other areas discussed in this ruling, the monitoring obligation in the statute is specific rather than general. Federal regulations require states to prioritize mediation in their monitoring activities. *See* 34 C.F.R. § 300.600(d)(2). Because the state does not do this, it is out of compliance with federal law.

B.

The next two issues relate to areas where the state is conducting some analysis, but the design of its processes prevents it from effectively monitoring school districts.

Targets. Another flaw in the state's monitoring system relates to the performance targets it has set for school districts. An adequate monitoring system requires adequate targets. If targets are too modest, states may fail to identify districts that are falling short on their obligation to provide an adequate education to disabled children.

While the IDEA grants states discretion in setting targets, that discretion is not limitless. For example, the IDEA requires states to design a state performance plan, in which states report to the federal government their progress in implementing the IDEA in school districts, using metrics and calculation methodologies that the Department of Education selects. States are given the discretion to set targets for those metrics, but the regulations specify that those targets must be "measurable and rigorous" across all priority areas. *See* 34 C.F.R. § 300.601(a)(3). The state uses many of the same metrics and targets both for monitoring districts and for the state performance plan. Dkt. No. 2478-1 at 6–7. The targets are set on a cyclical basis, and the state will revisit its targets in the coming months.

The problems with the state's chosen targets are clearest in the area of least restrictive environment, which is one of the IDEA's "priority areas" for monitoring. 20 U.S.C. § 1416(a)(3)(A); 34 C.F.R. § 300.600(d)(1). As discussed previously, the state uses three metrics to assess whether districts are placing school-aged disabled children in the least restrictive environment (placement in general education classes for greater than 80% of the day, for less

than 40% of the day, and in separate schools). And it has two metrics for preschool-aged children (placement in regular early childhood programs and in separate programs). In 2016, all five targets for these metrics were below statewide performance. Dkt. No. 2469 at 23, 26–27. And in many instances, statewide performance already fell significantly behind national performance. Thus, even though the data suggest that California performs poorly relative to the rest of the country at placing disabled students in general education placements, the state has set targets even below its already-poor statewide performance levels. This is particularly concerning in an area like least restrictive environment, because the state has elsewhere emphasized its strong belief that improving placements for disabled students will also improve outcomes for those students. *See, e.g.*, Dkt. No. 2506 at 62–63. Setting low targets increases the likelihood that districts will escape selection for targeted or intensive monitoring – and that districts will be classified by the state as compliant with the IDEA – even if they fall significantly short in placing disabled children in general education classes “to the maximum extent appropriate.” 20 U.S.C. § 1412(a)(5)(A).

The state’s targets for “child find” may be another example of flawed targets. The IDEA requires states to ensure that all children with disabilities, including homeless children and children attending private schools, are “identified, located, and evaluated.” 20 U.S.C. § 1412(a)(3)(A). Similar to least restrictive environment, the IDEA instructs states to prioritize child find in their monitoring activities. 34 C.F.R. § 300.600(d)(2). The state considers districts’ identification rates – the number of children identified as needing special education services over the total number of children served by the district. Districts that are more than two standard deviations from the statewide identification rate are selected for monitoring. But California’s identification rate is only 11.7% as compared to 13.2% nationally. Dkt. No. 2469 at 31. And in the 2017–2018 monitoring cycle, setting the target two standard deviations below California’s low mean resulted in a system that selected only those 38 districts (out of 1,296 districts) with an

identification rate below 3.6%.⁵ Dkt. No. 2455-1 at 35.

To be clear, in some instances it may be necessary to set modest targets to encourage incremental but stable progress. The state has explained that it follows a model of “continuous improvement,” which involves setting a baseline and encouraging development from that baseline. Dkt. No. 2478-1 at 3, 5. And for purposes of these court proceedings, the question is not whether the state has developed the best method for encouraging improvement (a question best left for the policymakers), but only whether the state’s system is capable of identifying districts that are out of compliance with federal law. But the monitor and plaintiffs have raised a genuine concern that the state’s targets are so low that they significantly impede its ability to identify districts experiencing serious performance issues.

The state first argued in its written submissions that all of its targets were adequate because they were the result of an inclusive process that considered input from various stakeholders. Dkt. No. 2478-1 at 5–7. That argument is unsatisfactory – the purpose of these court proceedings (and of the IDEA) is to ensure that the state has an adequate monitoring and enforcement system in place. Dkt. No. 2387 at 1. If the state’s performance targets are legally deficient – that is, if they prevent the state from adequately monitoring school districts – they cannot be defended by asserting that they were the result of a process that included many stakeholders. During the hearings, however, the policymakers agreed that, at minimum, the least restrictive environment targets are not adequate. They stated that these are not the targets they would have chosen had they been in charge at the time they were set. Dkt. No. 2507 at 101–03. The policymakers further confirmed that performance targets will be revisited in the coming

⁵ The plaintiffs and monitor also argue that the state is out of compliance because it only assesses districts with respect to child find for students ages 6–21. The state is interested in developing a measure for child find applicable to preschool-aged children, but finding an appropriate method has proven elusive. Because there’s no universal preschool in California, the state cannot assess how well districts are doing at identifying children with disabilities, because it lacks adequate information about the total number of children from which disabled students are being identified. The state has explained that there are projections available to estimate numbers of preschool-aged children based on many different variables, but those projections provide estimates by county (rather than by school district) and are not sufficiently reliable for use in the state’s data analysis activities. Dkt. No. 2507 at 132–38. This explanation is adequate.

months, and that they plan to adopt a different approach. Dkt. No. 2507 at 171.

The plaintiffs and monitor raised concerns with other targets set by the state. In response, at the hearings the policymakers often responded with arguments and statistics that were not raised in the state's written submissions, for example, concerning the number of districts that were selected using the relevant targets. Given that the state plans to revisit its performance targets so soon, there's no need to perform a target-by-target analysis. At this stage, what matters is that the record suggests that the state's targets do not enable it to adequately identify school districts for further monitoring. The state's new targets must be crafted with its monitoring obligations in mind, and it must do a far better job of explaining and justifying its targets during the renewed Phase 2.⁶

Comprehensive review. The next set of issues relate to the state's method for selecting districts for comprehensive review, the state's intensive monitoring activity. One might object that comprehensive review should receive less scrutiny for purposes of these proceedings, because the state's monitoring system is tiered. As a result, all of the school districts that are under serious consideration for comprehensive review have already been selected for one or more targeted monitoring activities. It is clear, however, from the state's written submissions and policymakers' testimony during the evidentiary hearings that there is a significant difference between the quality and depth of the state's monitoring activities, as well as the resources available to school districts, during targeted review versus intensive review. A closer look is therefore needed to determine whether the state's data analysis activities adequately sort districts between the targeted monitoring activities and comprehensive review.

This look reveals two major flaws in the state's current approach. The first is that the state appears to select the wrong districts. Under the state's formula, it appears that better-performing districts will often be far more likely to get selected for comprehensive review than

⁶ The state need not, however, present further information about its targets for the IDEA's timeliness requirements or for its risk-ratio for purposes of disproportionality monitoring, assuming those targets remain the same. As discussed later in this order, these targets do not appear to present any cause for concern and the state's explanations are adequate.

worse-performing districts, and the state has not provided an adequate explanation for the discrepancy. Second, the state's formula appears not to select enough districts for comprehensive review. The state's chosen metrics suggest that more districts need close scrutiny than currently receive it.

The state uses a point system to select districts for comprehensive review. For each metric, a district receives 1–4 points. Districts receive points for meeting the target, and more points for improving their performance from the prior year. For example, recall that the state evaluates districts' efforts to place disabled children in the least restrictive environment in part by calculating the percentage of disabled students in general education classes for greater than 80% of the day. The target for this metric is 51.2%. A district receives a "1" if it fails to meet the target and declines in performance from the previous year (for example, if the district places 45% of disabled children in general education classrooms for more than 80% of the day, but it placed 48% in the prior year). A district receives a "2" if it fails to meet the target but either improved or maintained its performance from the prior year. A district receives a "3" if it meets or exceeds the 51.2% target but declined in performance from the prior year. Finally, a district receives a "4" if it meets or exceeds the 51.2% target and either improves or maintains its performance from the prior year. Dkt. No. 2478-1 at 27.

Recall further that for the Dashboard metrics, school districts receive a color based on a grid that combines information about the district's current performance and any improvements or regressions from the prior year. In scoring districts for comprehensive review, districts receive a "1" for a red, a "2" for an orange, a "3" for a yellow, and a "4" for a "blue" or "green."

After the districts are scored, the state calculates a cut score and then selects all districts that fall below it for comprehensive monitoring. The cut score is expressed as a percentage of the total number of points that were available to that district, for all metrics for which the district was actually scored. According to the testimony of the policymakers, in the last monitoring cycle the state originally set the cut score at 65%, meaning that districts scoring less than 65% of all points available to that district would be selected for comprehensive review. But then, after the state ran

the numbers, it concluded that it did not have, or did not wish to spend, the resources necessary to put all those districts through comprehensive monitoring. So it reduced the cut score to 62%. Dkt. No. 2506 at 108–09, 153.

There are a number of problems with this system for selecting districts for comprehensive review. First, two features of the scoring methodology combine to produce odd results: (i) in awarding points, the state’s formula considers only whether a target is met or missed, without considering the extent by which a district missed or exceeded the target, meaning that near misses and wide misses receive the same number of points; and (ii) the state’s formula heavily prioritizes year-to-year improvements, even if the improvements are quite minor, meaning that minor fluctuations in performance can produce major differences in scoring. The court monitor identified several examples of how these two features influence districts’ scores. As discussed above, the state has set a target of 51.2% with respect to the placement of disabled students in general education classrooms for greater than 80% of the day.

In the last round of selection for comprehensive review, one particular district placed 62.8% of disabled students in general education classrooms for greater than 80% of the school day. This rate represented a 0.1% improvement from the prior year’s rate of 62.7%. The state awarded four points to that district. Another district had a rate of 92%, which represented a regression of 0.1% from the prior year’s rate of 92.1%. The state gave that district only three points, although it appears to have performed significantly better at placing students in general education classrooms, and the year-to-year differences for both are small. Dkt. No. 2469 at 60. In a different area, the state considers the number of times that it determined that a district violated federal law after investigating a parent’s complaint. One district went from 101 to 90 noncompliances and received a “2.” Dkt. No. 2506 at 106. By contrast, 17 districts increased from one noncompliance in the previous year to two in the current year and received a “1.”⁷

⁷ The state’s reasons for relying on annual changes rather than improvements over a more sustained period remain somewhat unclear. The Morgan Hill Concerned Parents Association makes a colorable argument in its amicus brief that comparing one year to the prior year does not provide the state with sufficient information to draw a reliable conclusion about the district’s

Another problem is that the state's formula is unweighted. It does not prioritize the metrics most closely related to the IDEA's guarantee of an appropriate public education for disabled students. For example, under the state's formula, school districts can receive up to four points merely for having high numbers of disabled students take statewide tests (regardless of how they perform). But districts can receive no more than four points for meeting actual performance targets for those assessments. Dkt. No. 2455-1 at 57. To use another example, districts can receive no more than two points for meeting five out of six preschool assessment targets, and for improving or maintaining performance from the prior year. But districts can also receive two points merely for submitting between 50% and 69% of all required reports to the state in a timely manner. Dkt. No. 2455-1 at 58, 62. The selection formula also entirely omits districts' performance on child find, although the IDEA singles out child find as a priority area for monitoring. 20 U.S.C. §§ 1412(a)(3)(A), 1416(a)(3)(B); 34 C.F.R. § 300.600(d)(2).

These problems appear to result in the selection of the wrong districts for comprehensive review. To assist the Court in evaluating the state's selection methodology, the court monitor developed a competing methodology and applied it to the same set of school districts evaluated by the state. The monitor's approach considered districts' relative distance from the target (rather than an all-or-nothing score), did not emphasize annual improvements, and focused only on four metrics that are particularly tied to disabled students' outcomes—performance on English and Language Arts and Math assessments, suspension rates, and placement of school-aged children in general education classes. Dkt. No. 2510 at 2. The monitor also expanded the analysis to include two more least restrictive environment metrics for school-aged children and preschool placements in regular early-childhood programs. Dkt. No. 2510 at 4. State policymakers agreed that the monitor chose metrics that are most closely tied to the provision of a free and appropriate education. Dkt. No. 2518 at 27. Finally, the monitor chose a cut score that would select roughly

trends in performance, as opposed to natural fluctuations in the data. Dkt. No. 2485 at 14; Dkt. No. 2485-1 at 3. The state did not offer a clear response in its written submissions or during the evidentiary hearings.

the same percentage of districts for monitoring that the state itself selected, to provide a fair point for comparison.

The monitor's method, when limited to four metrics, would select none of the districts selected by the state. What's more, all of the districts selected using this method scored 70% or more under the state's system—safely above the state's 62% cut-off for monitoring—or were not scored at all. Dkt. No. 2516 at 4. When the analysis is limited to districts with at least 31 students with disabilities, six districts selected by the state would be selected under the monitor's approach, but the majority of districts selected under the monitor's approach scored 70% or more under the state's approach. Expanding the analysis to seven metrics produces similar results. Dkt. No. 2516 at 5–6. The monitor's analysis suggests that the state's design choices significantly affect which districts are selected for comprehensive review, and that the districts selected are not likely the ones most in need of intensive intervention.

The data also suggest that not enough districts are selected. In all, the state scored 1,330 school districts for monitoring.⁸ Dkt. No. 2516 at 2. The state selected only 34 districts for comprehensive review. At the same time, excluding the first-year charters and very small school districts that the state simply ignored, the state classified only 48 districts as in compliance with federal law. Dkt. No. 2516 at 9. Perhaps these numbers would not provide much cause for concern if the data suggested that a significant proportion of districts were found out of compliance only in areas that are not as closely tied to the provision of an appropriate education. But the state's metrics further suggest that this is not the case. For example, 68 districts received a “red” on all three Dashboard metrics for suspension and performance on English and Language Arts and Math assessments, but only seven were selected for comprehensive review.⁹ (37 of those districts received a score of 70% or above under the state's selection formula).

⁸ This number excludes first-year charter schools. The state did not have prior-year performance data for these schools and therefore did not consider them for selection for comprehensive review.

⁹ This number also excludes first-year charter schools.

Policymakers testified that 83% of school districts were not meeting targets for the state’s English and Language Arts assessment, and 85% were not meeting targets for the state’s Math assessment. Dkt. No. 2506 at 192.

Furthermore, the state’s cut score does not appear to be tethered to a clear theory of compliance. In fact, it appears somewhat arbitrary. The state did not explain in its written submissions why a score of 62% (as opposed to a score of 52% or 72%) sufficiently identifies those districts urgently in need of intensive monitoring. And, as mentioned earlier, at the evidentiary hearings the policymakers suggested that the cut score was set, and then later adjusted, based on resource constraints. Dkt. No. 2506 at 153. Although it is inevitable that resource constraints and other practical considerations will play some role in the number of districts chosen for monitoring in a given cycle, it cannot, given the IDEA’s legal requirements, be the primary driver of the decision about where to set the cut score.

All these flaws in the state’s system for selecting districts for comprehensive review combine to put it in violation of the IDEA. The system does not result in meaningful choices about which districts, and how many districts, should receive intensive intervention.¹⁰ This is not to suggest that the state may never consider improvement in selecting school districts to monitor, or that the state may never use all-or-nothing scoring for an individual metric, if the targets for that metric are sufficiently demanding. Nor would the state be required to adopt the competing methodology designed by the court monitor. But the current system is so defective that it is not

¹⁰ As mentioned earlier, it’s worth noting that the state’s proposed selection methodology for its intensive preschool monitoring activity bears close similarities to the methodology used for comprehensive review. For preschool review, the state proposes to use an unweighted formula with all-or-nothing scoring (although the formula would not consider year-to-year performance). Districts would receive points across ten different metrics, and districts scoring “6” or more would be selected for review. (Confusingly, for comprehensive review, better-performing districts receive more points, while for preschool review, better-performing districts would receive fewer points). The cut score of “6” was set with resource constraints in mind and does not appear to be tied to a clear theory of statutory compliance. Given the similarities between the two methodologies, it seems reasonably likely that the same problems identified for comprehensive review would extend to preschool review as well.

merely inadvisable from a policy standpoint; it is legally inadequate.¹¹

III. FURTHER ISSUES

The issues described in Section II put the state out of compliance with federal law and prevent it from proceeding to Phase 3. The plaintiffs and monitor contend there are still other flaws in the state's procedures for identifying school districts for intervention. But in contrast to the issues discussed in Section II, the alleged deficiencies discussed in this section can all be placed in one of two categories: (i) the record does not justify a conclusion that the flaw exists in the first instance; or (ii) the flaw may exist, and perhaps it compounds the problems discussed in the previous section, but in the overall scheme, it does not appear significant enough to warrant a finding that the state is out of compliance with its statutory monitoring obligations. (That is, assuming the problems discussed in Section II will be adequately addressed.) Therefore, absent some significant revelation, or some significant change in the way the state conducts itself in these areas, the Court will not entertain further discussion on them during the renewed Phase 2.

A.

Disaggregation. The monitor and plaintiffs argue that the state is not in compliance across many metrics because, except as necessary for its disproportionality analyses, the state does not disaggregate its data. The state's monitoring system is designed to address outcomes for disabled students in general, and the state does not separately examine outcomes for disabled students who are also members of vulnerable subgroups, such as students in foster care, homeless students, English-language learners and students eligible for free and reduced-price lunch. The state also generally does not disaggregate its data by disability type.

The monitor's concern is that, if these vulnerable groups are small enough, the aggregate

¹¹ In addition to the issues discussed above, two issues identified during Phase 1 remain outstanding. During Phase 1, the state was found out of compliance for failing to collect data related to districts' implementation of their students' IEPs and the use of restraints and seclusion. Dkt. No. 2428 at 13–18, 26–28. The state is currently in the process of designing data collection protocols to address both issues. Because the state will need to accommodate these areas within all levels of its monitoring system, these issues remain relevant here, although they will be addressed in greater depth at a later date.

data will mask important information specific to these groups, so that a district that performs well with respect to disabled students as a whole will not be selected for monitoring even if it may fail to provide adequate services to, for example, disabled students that are in foster care.

This, however, is a policy issue rather than a legal one. Neither the monitor nor the plaintiffs have identified a legal basis for requiring the state to analyze these data, and the monitor stated at the hearing that he was unaware of any other state that disaggregates data in this manner. Dkt. No. 2506 at 169. By comparison, a separate set of statutory provisions, regulations, and guidance documents issued by the Department of Education govern a state's obligation to assess disproportionality for race and ethnicity, and even this form of analysis is limited to three discrete areas: discipline, identification, and placement.

Nor does it appear the failure to disaggregate as recommended by the monitor is significantly interfering with the state's ability to monitor school districts. To assess this issue, the monitor chose three metrics (two for least restrictive environment, and one for suspension) and applied them to a disaggregated sample of data. The monitor concluded that the disaggregation would result in the selection of districts for targeted reviews that were not previously selected by the state. In response, state officials observed that, assuming the validity of this approach, it would result only in incremental changes in the number of districts selected for monitoring. The state's "error rate," – meaning the number of school districts that were not selected for monitoring but should have selected under the monitor's approach, divided by the total number of school districts not selected – would range from somewhere between less than 1.0% to 3.6% for each metric. Dkt. No. 2518 at 81–82. Moreover, all of those districts are in some form of monitoring (and so none have been determined to "meet requirements"). Several are receiving assistance through the state's other accountability programs for poor outcomes related to foster children, regardless of disability.

Discipline. The state uses the Dashboard to assess district's discipline practices. The state considers only suspensions for longer than ten days. The plaintiffs and monitor argue that the state is out of compliance, because it does not consider multiple suspensions, and so a student

who is suspended once for one event is counted once, and a student suspended four times for four events in a year is also counted once. The state also does not consider the cumulative length of time a student is suspended. Dkt. No. 2469 at 12.

But neither the plaintiffs nor monitor have adequately explained how these issues significantly hinder the state's assessment of school discipline. It is not known, for example, whether one would expect to see many districts with many short suspensions but few lengthy suspensions. It is similarly unknown whether focusing on the number of suspensions per student would enhance the state's ability to red-flag districts. Perhaps one would ordinarily expect suspensions to be concentrated among smaller groups of students, and so focusing on students rather than suspension events would not add much. At this point, it is only apparent that the state does not use the most granular metric possible. It is not apparent that this practice materially affects the state's ability to monitor discipline practices.

Placements close to home. As part of the least restrictive environment principle, the IDEA requires districts to ensure that educational placements are as close as possible to the child's home, and in the school the child would attend if he or she was not disabled, "unless the IEP . . . requires some other arrangement." 34 C.F.R. § 300.116(b)(3), (c). At Phase 1, the state was found in compliance because it collected sufficient information to permit it to monitor district performance in this area, including information about students' "districts of residence" and "districts of service." *See* Dkt. No. 2435 at 175; Dkt. No. 2428 at 36–37.

The plaintiffs and monitor believe that the state should be ruled out of compliance because it does not do any data analysis for this requirement, or more generally into patterns of placement. State policymakers explained that this principle does not translate well to the state of California, because California is a "school choice" state, and so officials cannot presume that a child would attend the school closest to home if he or she were not disabled.

As applied to a state like California, the regulations are best understood as enshrining a general principle that districts must attempt to avoid concentrating disabled children in remote placements if more desirable placements are available. But it is not clear that data analysis

regarding placement is the best way to guard against this, for the reasons explained by the state. Perhaps analysis of IEP implementation – an issue on which the Court has ruled the state must collect data – would assist in this area. Or perhaps it’s an issue that should be considered only when the state is actually monitoring school districts (as opposed to when the state is assessing districts for different types of monitoring). In any event, the Court cannot conclude on this record that the state’s decision not to analyze “close to home” data when selecting districts for monitoring significantly interferes with its ability to meet its federal statutory obligations.

Complaints. The IDEA gives parents two mechanisms for enforcing their rights. First, it allows parents and advocates to file complaints with the state if they believe that a school district is not complying with federal or state special education law. 34 C.F.R. §§ 300.151–300.153. The state must carry out an investigation, if necessary, and resolve the complaint within 60 days. 34 C.F.R. § 300.152(a). Second, the IDEA permits parents or school districts to file due process complaints relating to the identification, evaluation, or educational placement of a disabled child, or the provision of a free and appropriate public education to the child. 34 C.F.R. § 300.507(a)(1). The parent or school district is entitled to a hearing before an administrative officer to resolve the issues raised in the complaint. 34 C.F.R. § 300.511(a), (c). A party dissatisfied with the officer’s decision can file a lawsuit in state court or federal district court. 34 C.F.R. § 300.516(a).

To evaluate district performance in this area, the state counts the number of instances in which a school district is determined to be out of compliance with federal law, through either process. These numbers do not lead to selection for any targeted review but are considered in selecting districts for comprehensive review. Districts receive one point for having more than one documented noncompliance, and another point if the number has decreased from the prior year. Districts with zero noncompliances receive four points. Dkt. No. 2455-1 at 60.

The plaintiffs and monitor object to the state’s use of a simple count rather than a per capita rate (the number of noncompliances divided by the total number of students within a district). Under the state’s method, two districts with three noncompliances each will be treated

the same, although three might be a troubling number for a very small district but less concerning for a very large district. The state responds that the target for this metric treats any noncompliance as giving significant cause for concern, and so using a more specific metric would not add materially to its analysis. In other words, the target for this metric is zero: districts with any noncompliances receive the lowest scores.

To be sure, there are problems with the state's explanation. The state does not merely rely on whether a district met or missed the target; it also considers whether the district has gotten better or gotten worse from the prior year. It's hard to see how the state could meaningfully determine whether the district is improving or regressing from raw totals alone and without any sense of the size of the school district. A district's population may fluctuate from year to year, and so a district's number of noncompliances could conceivably increase as the result of population changes, rather than as a result of poor district performance.

That said, the record does not demonstrate that this issue, which represents one point in a complicated selection formula, materially affects the state's ability to monitor school districts. And because substantial changes must already be made to the state's selection methodology for comprehensive review, this issue is likely to become even less important by the time of the renewed Phase 2 proceedings.

Preschool achievement. The state administers assessments for both school-aged children and preschool-aged children. For preschoolers, the state uses the "Desired Results Developmental Profile," which measures social development, knowledge acquisition, and the use of appropriate behaviors to meet needs. Dkt. No. 2455-1 at 25. The state has established six targets for disabled students taking these assessments. However, there is no targeted review for preschool achievement. Poor performance in preschool assessments is one factor in the state's proposed preschool review, but a minor one. A district that misses all six assessment targets would receive one point, and a district needs a total of six points for selection. Performance on preschool assessments also comprises only one of many different selection metrics for comprehensive review. As a result, a school district that performs poorly on its preschool

assessments will not receive any monitoring, unless it also performs poorly across other dimensions. The plaintiffs and monitor believe that poor performance on preschool assessments should receive a separate targeted review.

The state explained that it has made a policy determination to focus its efforts on placing more preschoolers in regular early childhood programs rather than separate programs. Dkt. No. 2506 at 204. State officials testified that placements in general education settings (where expectations are often higher) are associated with better performance on assessments. In addition, the state's data establishes that all districts that missed one or more preschool assessment targets have been selected for at least one form of monitoring. Dkt. No. 2503 at 1. Of those districts, half are monitored for inadequate placements. Further, because all of these districts are in some form of monitoring, none has received a positive annual compliance determination from the state. Given this record, the Court will not second-guess the state's policy determination that for now, the best way to evaluate and monitor districts with respect to preschool-aged students is to focus on preschool placements.

Referrals. As discussed above, the state assesses school discipline by looking at districts' rates of suspensions greater than ten days. The state does not collect data regarding districts' referrals of disabled students to law enforcement. During Phase 1, the Court noted the importance of this data, because disabled students are often referred to law enforcement at disproportionate rates. However, the state was not ruled out of compliance for failing to independently collect referral data, because the federal government routinely collects the data and makes it publicly available. Accordingly, the Court found at Phase 1 that the state had sufficient access to referral data and reserved for Phase 2 the question whether the state adequately analyzes this data in determining which districts to select for monitoring.

During Phase 2, the state took the position that it doesn't need to separately analyze referral data, because school districts' suspension rates are correlated with their rates of referrals to law enforcement, such that suspension data alone is sufficient to red-flag districts who may be overusing referrals to law enforcement. The record supports this. Although the referral data that

is publicly available is of poor quality, the monitor's supplemental analysis suggests a meaningful correlation between school districts' suspension rates and their referral rates. Dkt. No. 2510 at 10–12. The analysis also suggests that there are few districts with high referral rates but low suspension rates. Dkt. No. 2510 at 12. While it remains conceivable that the state may need to show that during its actual monitoring activities it sufficiently disentangles the reasons that referrals to law enforcement occur, there is not an issue at this phase with respect to law enforcement referrals. As with preschool achievement, the Court will not second-guess the state's decision to use suspension practices as a proxy for referral practices on this record.

Parent input. The state also analyzes parent input as part of its monitoring processes. The state asks parents to respond to the question: “Did the school district facilitate parent involvement as a means of improving services and results for your child?” Dkt. No. 2455-1 at 30. To analyze this data, the state records the percentage of “yes” responses for each school district. Districts that fall below the target (92% of parents responding “yes”) are selected for performance indicator review and move closer to selection for comprehensive review. The plaintiffs and monitor believe that this target is unambitious because in 2016, state performance was at 99.4%. Dkt. No. 2469 at 40. But for many of the reasons discussed in the Court's Phase 1 order and at the Phase 1 hearings, data about parents' responses to this question do not appear to be particularly useful (nor does the Court have confidence that any sort of statewide parent survey data would be actionable), so any concerns about the target set by the state in connection with it will not give rise to a noncompliance finding.

B.

In addition to the above metric-level issues, the Court finds that two of the state's selection methods for monitoring activities present no cause for legal concern on this record:

Data identified noncompliance. The state's method of selecting districts for data identified noncompliance monitoring (the monitoring activity relevant to the IDEA's timeliness

requirements) appears to be adequate.¹² Because the targets require perfect performance, a single untimely IEP meeting or initial evaluation is sufficient to flag a district for monitoring. The plaintiffs note that the state currently does not monitor districts for compliance with the IDEA's deadlines for resolution sessions, which are meetings that districts must convene between parents and members of the disabled student's IEP team after the parent files a due process complaint against the district. The resolution session must be held within 15 days after the district receives notice of the complaint. 20 U.S.C. § 1415(f)(1)(B)(i)(I); 34 C.F.R. § 300.510(a)(1). But going forward, the state plans to include timeliness in convening resolution sessions as a metric for targeted review. Dkt. No. 2506 at 15; Dkt. No. 2455-1 at 8 n.17.

Disproportionality. The state's selection formulas for its three disproportionality reviews are, once the state completes a couple of simple fixes, adequate. These are the state's monitoring activities for districts where students appear to be disproportionality represented among those receiving special education services based on their racial or ethnic group. Federal regulations require the state to use a specific formula to assess disproportionality, although states retain discretion to set an appropriate target for determining when a school's numbers suggest that it has become "significantly disproportionate." 34 C.F.R. §§ 300.646, 300.647.

One concern is that while the state appears to have set an appropriate target, its current formula for assessing disproportionality in discipline and placement is incorrect. The state's formula compares the risk of disciplinary action or restrictive placements for disabled students of a particular race or ethnicity to the risks faced by students of other races and ethnicities who are in general education settings. Dkt. No. 2469 at 17. The United States Department of Education, however, has explained that the relevant comparator is the risk faced by disabled students of other races and ethnicities. The state has conceded the error and plans to put the correct formula

¹² The state currently assesses compliance with the IDEA's deadlines for (1) evaluating students to determine eligibility for special education services, 34 C.F.R. § 300.301(c)(2)(i); (2) putting IEPs in place for toddlers transitioning to preschool, 34 C.F.R. § 300.124(b); (3) performing required reviews of IEPs, 34 C.F.R. § 300.324(b)(1)(i); (4) conducting required reevaluations of students with IEPs, 34 C.F.R. § 300.303(b)(2).

in place for the 2019–2020 school year. Dkt. No. 2507 at 68. Given this commitment, and the absence of any other outstanding issues with respect to the targeted disproportionality review, this issue does not provide a basis for finding the state out of compliance with its statutory obligations.

The state is also in compliance with respect to significant disproportionality review, the intensive form of monitoring. The state selects all districts that are disproportionate (as defined by the state’s chosen threshold) for three years for intensive monitoring. Federal law permits a state to exempt districts from monitoring if it determines, in its discretion, that a district is making “reasonable progress” towards resolving its issues. 34 C.F.R. § 300.647(d)(2). The state does not exercise this discretion but selects for monitoring all districts that exceed the threshold. Because the state’s formula selects more districts for monitoring than federal law requires, the state is currently in compliance. If the state adopts a “reasonable progress” definition in the future, and if the plaintiffs believe that the definition is so deficient that it would significantly interfere with the state’s ability to identify and select school districts for disproportionality monitoring, they may bring that to the Court’s attention. But absent any such development, there is no need for further inquiry during this phase.

Private school monitoring. During the Phase 1 hearings, policymakers testified that the state monitors private schools that educate students with disabilities who are placed there by virtue of their IEPs. These schools must be certified by the Special Education Division, and although the state collects data from these schools, they are all monitored on a cyclical basis, irrespective of whether the data suggests they need it. Dkt. No. 2518 at 94–98. Because the state monitors all private schools, this area presents no issues for purposes of Phase 2, although Phase 3 will involve further inquiry into the content and quality of the state’s monitoring activities.

Effectiveness of monitoring. The IDEA states that the “primary focus” of any state’s monitoring system must be on “[i]mproving educational results and functional outcomes for all children with disabilities,” and ensuring that school districts comply with the law, including

“those requirements that are most closely related to improving educational results for children with disabilities.” 34 C.F.R. § 300.600(b). This focus logically requires the state to conduct some assessment of the effectiveness of its monitoring activities – in effect, to monitor its monitoring system. The monitor and the plaintiffs argue that the state is out of compliance because it did not present any information in this phase about how it does this.

It’s true that, to get out from under the consent decree, the state must show that it has institutionalized a process for self-examination, to ensure that its monitoring actually improves student outcomes and that its data analysis activities do not become detached from realities on the ground. The deficiencies identified in this order underscore the importance of such a system. However, it makes little sense to require the state to present evidence about its self-evaluation process at this stage, given that it is already engaging in self-reflection as part of these proceedings, and already considering substantial changes to its monitoring system. Accordingly, the state need not revisit this requirement until Phase 4.

IV. NEXT STEPS

Given the concerns discussed in this ruling, the state must repeat Phase 2. In the next set of written submissions and at the next set of evidentiary hearings, the state must explain the changes that it has made to address the concerns identified in this order. The sequence of written submissions will remain the same, although the parties are urged to include data analyses where appropriate.

A case management conference is scheduled for Wednesday, August 28, 2019 at 9:30 a.m. to discuss the next phase and the state’s proposals for addressing IEP implementation. A case management statement is due seven days prior. The parties should include a proposed schedule for the written submissions and evidentiary hearings.

The parties may also, if they choose, file any materials with the case management statement that would assist the Court in facilitating a discussion between the parties regarding the state’s proposals for integrating IEP implementation into its monitoring protocols. If the parties no longer need to hold such a discussion with the Court (or if they wish to defer the matter to the

next set of evidentiary hearings), then the policymakers do not need to attend the case management conference, although they are welcome to do so, either in person or by phone.

IT IS SO ORDERED.

Dated: July 5, 2019

A handwritten signature in black ink, appearing to read 'V. Chhabria', is written above a horizontal line.

VINCE CHHABRIA
United States District Judge